1 ## TITLE OF THE INVENTION

2 "VIDEO CODING BY ADAPTIVELY CONTROLLING THE

3 INTERVAL BETWEEN SUCCESSIVE PREDICTIVE-CODED

4 FRAMES ACCORDING TO MAGNITUDE OF MOTION"

5 ## BACKGROUND OF THE INVENTION

6 ### Field of the Invention

7 The present invention relates to a video coding apparatus and

8 method for encoding moving pictures in a compressed format according

9 to an international standard such as ISO/IEC 13818-2, known as

10 MPEG-2.

11 ### Description of the Related Art

12 The MPEG-2 standard defines three picture types: intra-coded

13 pictures (I-pictures), predictive coded pictures (P-pictures) and bi-

14 directionally predicted pictures (B-pictures). I-pictures are coded in

15 such a way that they can be decoded without knowing anything about

16 other pictures in a video sequence. The first picture in a group of

17 pictures is always an I-picture and provides key information for pictures

18 that follow. P-pictures are coded (i.e., forward predictive coded) by

19 using information from a reference picture displayed earlier, which may

20 be either an I-picture or a P-picture. B-pictures also use information

21 from pictures displayed earlier and from pictures coming in the future

22 (i.e., forward-and-backward predictive coded). These three picture

23 types cyclically occur in a predetermined pattern. According to the

24 current practice, I-pictures occur at intervals represented by an integer

25 N and the interval between an I-picture and a P-picture is represented

1 by an integer M. Since these integers are of fixed value, the video

2 sequence is dynamically controlled so that they are maintained constant.

3   As shown in Fig. 1A, when M = 1, an I-picture is followed by a

4 sequence of P-pictures. Each P-picture is coded by using information

5 from a picture that immediately precedes it. For M = 2 (Fig. 1B), the

6 interval between an I-picture and a P-picture is equal to "2" and a B-

7 picture comes in between. In this case, each P-picture is coded by using

8 information from a picture preceding it by two-picture interval and

9 each B-picture is coded using information from two pictures, one

10 immediately preceding it and the other immediately succeeding it. For

11 M = 3 (Fig. 1C), the interval between an I-picture and a P-picture is

12 equal to "3" and two B-pictures are used to fill in the interval. In this

13 case, each P-picture is coded by using information from a picture

14 preceding it by three-picture interval, and one of the two B-pictures is

15 coded using information from an immediately preceding picture and

16 from a future picture that comes two-picture intervals following it, and

17 the other of the two B-pictures is coded using information from a

18 previous picture that precedes it by two-picture intervals and from a

19 picture immediately following it. Thus, for M ≥ 2, the number of B-

20 pictures that come in between non-B-pictures is equal to M − 1.

21   One reason for using the B-picture is to reduce the amount of

22 redundant video information inherently contained in the original frame.

23 For a given quantization scale, the use of B-pictures can reduce the

24 number of codes with which original pictures are encoded. Hence, the

25 picture quality can be improved for a given compression (coding) rate.

1 Another reason for using the B-picture is its tendency toward cancelling

2 an accumulated error that will result from continued prediction coding

3 processes that use information only from previous pictures of "parent

4 generations" which themselves were predicted from reference pictures of

5 "grandparent generations". Therefore, if unidirectional (forward)

6 predictive coding were exclusively used, predictive coded "generations"

7 would increase rapidly with time and quantization errors would

8 accumulate significantly. B-pictures present a solution to this problem.

9 Although the B-picture provides a benefit, the use of many B-

10 pictures (with the resultant increase in the M-value) is disadvantageous

11 for fast-moving pictures since it becomes difficult to search for motion

12 vectors within a range that is considered appropriate. Consider, for

13 example, an object moving at a constant velocity. Since the amount of

14 motions for each frame is constant, an increase in the M-value would

15 cause the moving object to proportionally increase its range of motions.

16 In order to precisely search for motion vectors, it would be necessary to

17 perform a vector search over a wide range that is variable in proportion

18 to the M-value.

19 One prior art approach involves setting a maximum value of

20 per-frame motions and then determining a range of motion vectors to be

21 searched for that is M times the maximum value. However, a

22 significant amount of hardware is necessary to implement this approach.

23 Although the hardware problem can be avoided by the use of an

24 algorithm that simplifies motion vector search, this would be only

25 achieved at the cost of search precision and a poor picture quality would

1 result.

2 Another prior art approach is disclosed in Japanese Laid-Open

3 Patent Application 9-294266. According to this technique, a

4 distribution of motion vectors and a differential value of inter-frame

5 predictions are detected. The M-value is increased according to the

6 detected distribution and is decreased according to the detected

7 differential value. Therefore, if a motion-vector search is being

8 performed on a current P-picture using M = 2 over a given range and

9 most of the motion vectors are found to exist in that given range, then the

10 M-value is incremented to 3 and a picture that is three frame intervals

11 future from the current P-picture is determined as the next P-picture.

12 Otherwise, the M value remains unchanged and a picture that is two

13 frame intervals future from the current P-picture is determined as the

14 next P-picture. If the detected differential value of inter-frame

15 predictions exceeds some threshold, the M-value is decremented to 1 and

16 a picture that is one frame interval future from the current P-picture is

17 determined as the next P-picture. However, it is established that, in

18 most cases, the distribution of motion vectors is isotropic about an

19 average vector and its spread (variance) varies depending on the

20 strength of auto-correlation of motions. Therefore, statistical data of

21 motion vectors cannot be estimated by the number of motion vectors

22 which exist in a search range and exceed a threshold value. If motion

23 vectors have a large mean value in the neighborhood of a threshold

24 within a given range that is considered sufficient for a search regardless

25 of their variance, the narrowing of the search range would cause a

1 significant degradation of picture quality. If the distribution of motion

2 vectors is used for making a decision for the adequacy of the search

3 range and if the algorithm for such decision is based solely on a motion

4 vector distribution approaching a zero vector point, a decision is likely

5 to be made in favor of the adequacy of the search range. When the

6 distribution immediately moves away from the zero vector point, it can

7 occur that the search range will be found to be insufficient. Therefore,

8 several frames would be taken to readjust the interval between

9 successive P-pictures. A delayed action will cause poor picture quality.

10 ## SUMMARY OF THE INVENTION

11 It is therefore an object of the present invention to provide a video

12 coding apparatus and method (algorithm) for reducing hardware scale

13 while enabling a wide range of motion vectors to be searched for.

14 In general terms, the video coding apparatus comprises

15 coding/decoding circuitry for providing motion-compensated inter-

16 frame prediction coding on input frames by using reference frames so

17 that the input frames are coded into an intra-frame coded picture, a

18 predictive coded picture or a bi-directionally predictive coded picture

19 and decoding the coded frames to produce reference frames. Decision

20 circuitry is provided for determining the magnitude of motion of the

21 input frames relative to the reference frames, determining the interval

22 between successive frames of the predictive coded pictures and

23 reordering the input frames according to the determined interval.

24 In specific terms, the video coding apparatus of this invention

25 comprises a first memory for storing a plurality of input frames, a second

1   memory for storing reference frames, motion vector detection circuitry

2   for detecting motion vectors in frames from said first memory relative to

3   reference frames selectively supplied from said second memory

4   according to a control signal, coding/decoding circuitry for providing

5   motion-compensated inter-frame prediction and coding on a frame

6   supplied from said first memory according to the detected motion

7   vectors and said control signal so that the frame is coded into an intra-

8   frame coded picture, a predictive coded picture or a bi-directionally

9   predictive coded picture and locally decoding the coded frame and

10 storing the decoded frame in said second memory as one of said

11 reference frames, and mean value calculation circuitry for calculating,

12 at frame intervals, a mean value of the detected motion vectors.

13 Decision circuitry determines an interval between successive frames of

14 said predictive coded picture according to the mean value, and

15 modifies the control signal according to the determined interval.

16   According to a further aspect, the present invention provides a

17 video coding method comprising the steps of providing motion-

18 compensated inter-frame prediction and coding on input frames by

19 using reference frames so that the input frames are coded into an intra-

20 frame coded picture, a predictive coded picture or a bi-directionally

21 predictive coded picture, decoding the coded frames to produce the

22 reference frames, determining the magnitude of motion of the input

23 frames relative to the reference frames, determining the interval between

24 successive frames of the predictive coded picture according to the

25 determined magnitude of motion and reordering the input frames

1 according to the determined interval.

2 ## BRIEF DESCRIPTION OF THE DRAWINGS

3 The present invention will be described in further detail with

4 reference to the accompanying drawings, in which:

5 Figs. 1A, 1B and 1C are illustrations of sequences of frames for

6 different M-values according to the MPEG-2 standard;

7 Fig. 2 is a block diagram of a video coding apparatus according

8 to one embodiment of the present invention;

9 Fig. 3 is a flowchart of the operation of the GOP structure

10 decision circuit according to one embodiment of the present invention;

11 Figs. 4A and 4B are diagrams for illustrating the relationships

12 between input (display) order and coding order;

13 Fig. 5 is a schematic illustration of the search ranges of the video

14 coding apparatus;

15 Fig. 6 is a block diagram of a practical form of the video coding

16 apparatus of the present invention;

17 Fig. 7 is a flowchart of the operation of the GOP structure

18 decision circuit according to a modified embodiment of the present

19 invention; and

20 Fig. 8 is a graphic representation of average motion vector, rate of

21 change of average motion vector and an M-value plotted versus frames.

22 ## DETAILED DESCRIPTION

23 Referring to Fig. 2, there is shown a video coding apparatus

24 according to the present invention. The coding apparatus is comprised

25 of an input frame memory 101 for receiving a plurality of video frames

1 supplied from an input terminal 150 for storage and outputting frames

2 in a coding order in which these output frames will be encoded. Each

3 of the stored frames is divided into a plurality of regions or

4 "macroblocks" and a coding process will be performed on each of the

5 macroblocks. The reordering of the frames in the input frame memory

6 101 is controlled by a GOP structure decision circuit 110 which

7 produces an M-value representing a GOP (group of pictures) structure.

8 A differential signal representing the error between a predicted frame

9 provided by an motion-compensated inter-frame predictor 104 and a

10 frame supplied from memory 101 is produced by a subtractor 105. This

11 prediction error is coded by an encoder 106 and supplied to an output

12 terminal 151.

13 The output of encoder 106 is further connected to a decoder 107 to

14 reconstruct the prediction error, which is combined in an adder 108

15 with the frame predicted by motion-compensated inter-frame predictor

16 to produce locally decoded frames. The locally decoded frames are

17 stored in a reference frame memory 102 as reference frames, which are

18 then selected and delivered to the motion-compensated inter-frame

19 predictor 104 and a motion vector searcher 103. The selection of the

20 reference frames is determined by the GOP structure decision circuit

21 110.

22 Motion vector searcher 103 receives input frames from the input

23 frame memory 101 and reference frames from the reference frame

24 memory 102 and makes a search through a range of input frames

25 determined by the GOP structure decision circuit 110 for detecting

1   motion vectors. If the video signal is interlaced, the motion vector

2   searcher 103 may be configured to make a search through each field of

3   input frames. Alternatively, each field of the frame may be divided into

4   a plurality of blocks of 16 pixels by 8 lines each. In this case, the motion

5   vector search may be provided for each of these blocks.

6   Motion-compensated inter-frame predictor 104 uses a reference

7   frame from the memory 102 to provide a motion-compensated inter-

8   frame prediction on an input frame from the memory 101 in accordance

9   with the output of the GOP structure decision circuit 110 and the

10   output of the motion vector searcher 103, so that the input frame is

11   coded by the encoder 106 as an I-picture, a P-picture or a B-picture

12   depending on the M-value determined by the decision circuit 110.

13   When intra-frame coding is performed, the motion-compensated inter-

14   frame predictor 104 produces no output signal. In this case, a frame or a

15   macroblock from the input frame memory 101 is passed through the

16   subtractor 105 without alterations and supplied to the encoder 106.

17   A mean value calculator 109 is connected to the motion vector

18   searcher 103 to calculate a mean value of motion vectors detected by the

19   searcher 103 from each macroblock and produces an average motion

20   vector for each macroblock. Since the motion vectors detected by the

21   searcher 103 are vectors in the forward direction (from previous to

22   current) as well as the backward (from future to current) direction, a

23   weighted mean value is calculated using one motion vector in a

24   macroblock.

Sub 0.   ~~In an alternative embodiment, the motion vector searcher 103~~

1   makes a decision as to whether or not intra-frame coding is appropriate

2   during a search through macroblocks of the input frame. If this is the

3   case, the motion vector searcher 103 provides no output to the mean

4   value calculator 109.

5   The GOP structure decision circuit 110 is configured to produce

6   a signal indicating whether the frame currently being encoded is an I-

7   picture, a P-picture or a B-picture and supplies picture type indication

8   to the motion vector searcher 103. The GOP structure decision circuit

9   110 proceeds to perform an M-value updating process according to the

10 flowchart of Fig. 3.

11   The routine begins with initialization of the M-value at step 200.

12 At decision step 201, the decision circuit 110 determines from the

13 current M-value whether or not the current frame is a predictive coded

14 picture (P-picture). If the decision is affirmative at step 201, the routine

15 proceeds to step 202 to receive a weighted mean value from the mean

16 value calculator 109 and determines, at step 203, whether or not the

17 current frame is a still picture. If the current frame is a still picture, the

18 routine proceeds to decision step 204 to determine if the current M-value

19 is smaller than a predefined maximum value. If the current M-value is

20 smaller than the maximum value, the decision circuit 110 increments M

21 by a prescribed amount (step 205).

22   If the current frame is not a still picture, the routine proceeds

23 from step 203 to decision step 206 to check to see if the current P-picture

24 is a fast moving picture. If the current frame is a fast moving picture,

25 the decision circuit 110 determines, at step 207, whether the current M-

1 value is greater than 1. If M is greater than 1, the M-value is

2 decremented by a prescribed value (step 208).

3   Following the execution of step 205 or 208, the decision circuit

4 110 proceeds to step 209 to control the frame memories 101, 102, the

5 motion vector searcher 103 and the motion-compensated inter-frame

6 predictor 104 according to the updated M-value, and then returns to

7 step 201 to repeat the M-value updating process. At decision step 201 of

8 each successive updating process, the decision circuit 110 determines the

9 picture type from the M-value updated in a previous process.

10   If each of the decisions made at steps 201, 204, 206 and 207 is

11 negative, the decision circuit 110 proceeds to step 209 to control the

12 memories 101, 102, searcher 103 and predictor 104 according to the

13 current M-value.

14   The starting value with which the M-value is initialized may be a

15 maximum value. At steps 205 and 208, the M-value may be varied with

16 a unit value of one.

17   The updating process of the GOP structure decision circuit 110

18 will be visualized by the following description with reference to Figs.

19 4A and 4B by assuming that the input memory 101 has the capacity of

20 storing as many frames as necessary to provide reordering when the M-

21 value is maximum.

22   In Fig. 4A, the M-value is successively decremented if the GOP

23 structure decision circuit 110 determines that a current frame is a fast

24 moving picture. If the initial M-value is 3, input frames are stored in

25 the input frame memory 101 in the order 1(B), 2(B), 3(I), 4(B), 5(B) and

1   6(P) for a period necessary for reordering. These stored frames are

2   reordered such that the third frame 3(I) comes first in the coding order

3   so that it can be intra-frame coded as an I-picture I3. The I-picture I3 is

4   followed by the first and second frames 1(B) and 2(B) so that they are

5   coded as B-pictures B1 and B2. The sixth frame 6(P) comes in the

6   fourth position so that it can be coded as a P-picture P6. The P-picture

7   P6 is followed by the fourth and fifth frames 4(B) and 5(B), which will

8   be coded as B-pictures B4 and B5. When the M-value is decremented to

9   "2", subsequent input frames are stored in the frame memory 101 in the

10 order 7(B), 8(P), 9(B) and 10(P) for a period necessary for reordering.

11 Since the frame 8(P) must precede the frame 7(B), these frames are

12 reordered and coded as a P-picture P8 and a B-picture B7. Likewise,

13 since the frame 10(P) must precede the frame 9(B), these frames are

14 reordered and coded as a P-picture P10 and a B-picture B9. When the

15 M-value is further decremented to "1", subsequent frames are stored in

16 the frame memory 101 in the order 11(P), 12(P) and 13(P). In the

17 illustrated example, since the output frames are delayed by two frames

18 with respect to the input frames, input frames 11(P), 12(P) and 13(P) are

19 stored for a two-frame interval and delivered without being reordered

20 and encoded into P-pictures P11, P12 and P13.

21   In Fig. 4B, the M-value is successively incremented if the GOP

22 structure decision circuit 110 determines that a current frame is a still

23 picture, starting from M = 1 in which input frames are stored in the

24 frame memory 101 in the order 14(P), 15(P) and 16(P) for a two-frame

25 interval and delivered without being reordered and encoded as frames

1   P14, P15 and P16. When the M-value is incremented to "2", frames

2   17(B) and 18(P) are reversed in order, and frames 19(B) and 20(P) are

3   reversed in order and delivered as P18, B17, P20 and B19. When the M-

4   value is incremented to "3", frame 23(P) comes earlier than frames 21(B)

5   and 22(B), and frame 26(P) comes earlier than frames 24(B) and 25(B).

6   All of these frames are delivered as P23, B21, B22, P26, B24 and B25.

7   Fig. 5 schematically illustrates search ranges of the motion vector

8   searcher 103 for different M-values in a one-dimensional scale (note that

9   the actual search ranges are two-dimensional). Assume that the motion

10 vector searcher 103 is making a search in a given direction in the range

11 between R_1 and R_2. Consider the relationships between a current frame

12 and reference frames (P-pictures) with M = 1, 2 and 3. As shown in Fig.

13 5, when the searcher 103 is making a search with M = 1, the increment

14 of frame interval by one frame will enlarge the search range by a factor

15 of 2 and the increment of frame interval by two frames will enlarge the

16 range by a factor of 3. Likewise, when the searcher 103 is making a

17 search with M = 2, the increment of frame interval by one frame will

18 enlarge the search range by a factor of 3/2. In this way, when there is a

19 fast moving object the searcher 103 can keep track of its motion vectors.

20   When fast moving objects are detected and hence it is difficult to

21 perform a wide-range vector search, the GOP structure decision circuit

22 110 decrements the M-value to shorten the frame interval between P-

23 pictures so that the performance of the motion-compensated inter-frame

24 predictor 104 is increased. In contrast, when still pictures are detected

25 and hence it is not necessary to perform a wide-range vector search, the

1   GOP structure decision circuit 110 increments the M-value to lengthen

2   the frame interval between P-pictures to increase intervening B-pictures

3   so that the overall coding efficiency of the apparatus is improved.

4   Fig. 6 shows a practical form of the video coding apparatus of the

5   present invention in which parts corresponding in significance to those

6   of Fig. 2 are marked with the same numerals and the description thereof

7   is omitted for simplicity. The encoder of Fig. 2 is replaced with a DCT

8   (discrete cosine transform) coder 600 and a quantizer 601 and the

9   decoder is replaced with a dequantizer 602 and an inverse DCT circuit

10   603. DCT coefficient data is quantized by the quantizer 601 and

11   supplied to a variable length coder 605 as well as to the dequantizer 602.

12   Variable length coder 605 performs run-length coding on the quantized

13   DCT coefficient by using the motion vector supplied from the motion

14   vector searcher 103.

15   The flowchart shown in Fig. 7 is used to operate the GOP

16   structure decision circuit of Fig. 6. The M-value is first initialised at

17   step 700 and the picture type is determined at step 701 from the M-

18   value. If a P-picture is detected, the GOP structure decision circuit 110

19   proceeds to step 702 to receive a mean value of horizontal motion vectors

20   (MVavex) and a mean value of vertical motion vectors (MVavey) from

21   the mean value calculator 109. At step 703, the decision circuit 110

22   compares the M-value with a reference value "1". If M = 1, flow

23   proceeds to step 704 to check to see if the following conditions are

24   simultaneously met:

25 |MVavex| is smaller than a threshold value Th1; and

1 |MVavey| is smaller than a threshold value Th2.

2 If the above-mentioned conditions are met, the decision circuit 110

3 proceeds to step 705 to calculate the rate of change of horizontal average

4 motion vector (ΔMVavex) and the rate of change of vertical average

5 motion vector (ΔMVavey) as follows.

6 ΔMVavex = MVavex (t) – MVavex (t – 1)

7 ΔMVavey = MVavey (t) – MVavey (t – 1)

8 where t represents the frame number.

9 GOP structure decision circuit 110 proceeds to step 706 to

10 determine whether the following conditions are simultaneously satisfied:

11 |ΔMVavex| is smaller than threshold Th3; and

12 |ΔMVavey| is smaller than threshold Th4.

13 If these conditions are simultaneously met, flow proceeds to step 707 to

14 increment the M-value by a predetermined amount.

15 If M is greater than 1, the decision circuit 110 proceeds from step

16 703 to step 708 to determine if one of the following conditions is met:

17 |MVavex| is equal to or greater than a threshold value Th5; or

18 |MVavey| is equal to or greater than a threshold value Th6.

19 If the decision at step 708 is affirmative, flow proceeds to step 709 to

20 decrement the M-value by a predetermined amount. At step 710, the

21 decision circuit 110 controls the frame memories 101, 102, the motion

22 vector searcher 103 and the motion-compensated inter-frame predictor

23 104 according to the updated M-value, and then returns to step 701 to

24 repeat the M-value updating process.

25 If each of the decisions made at steps 701, 704, 706 and 708 is

1 negative, the decision circuit 110 proceeds to step 710 to control the

2 memories 101, 102, searcher 103 and predictor 104 according to the

3 current M-value.

4 In order to evaluate the performance of the video coding

5 apparatus, fast moving pictures were experimentally used as input

6 frames and the incremental unit of steps 707 and 709 was set equal to 2

7 so that the M-value is switched between 1 and 3 when each of the

8 decisions at steps 706 and 708 is affirmative. Results of the experiment

9 are shown in Fig. 8, in which the horizontal average motion vector, the

10 rate of change of the horizontal average motion vector and the

11 corresponding M-value are plotted as a function of the number of

12 frames. It is seen that when a fast moving picture is detected, the M-

13 value is reset to 1. Since the time-varying rate of motion vectors is taken

14 into account by steps 705 and 706, the M-value is maintained at 1 when

15 the number of frames is 170, where the average motion vector is crossing

16 the zero level. In this way, the updating performance of the M-value is

17 optimized.

18 The average motion vectors MVavex and MVavey used by the

19 GOP structure decision circuit 110 could be altered in a number of

20 ways. By designating such vectors as |MV|, either one of the following

21 conversions can be used:

22 $\qquad |MV| = |MVavex| + |MVavey|$

23 $\qquad |MV| = MVavex^2 + MVavey^2$

24 $\qquad |MV| = $ Square root of $(MVavex^2 + MVavey^2)$

25 $\qquad |MV| = a * |MVavex| + b * |MVavey|$

1        $|MV| = $ Square root of $(a * MVavex^2 + b * MVavey^2)$

2 where a and b are constants.